

WHAT IS CLAIMED IS:

1. A method for predicting the interaction between proteins, comprising the steps of:
- (a) obtaining appearance frequency information of a designated domain combination selected from each of interacting and non-interacting sets of protein pairs, and storing the obtained appearance frequency information;
- (b) determining a probability equation applied to predict the interaction between two proteins using the stored appearance frequency information of the domain combination; and
- (c) obtaining an interaction probability value between the two proteins from the determined probability equation.

2. The prediction method as set forth in claim 1, wherein the step (a) includes the sub-steps of:

- (a-1) creating a weighted appearance frequency; and
- (a-2) creating an appearance probability (AP) matrix based on the weighted appearance frequency.

3. The prediction method as set forth in claim 1, wherein:
- the appearance frequency information of the domain combination is defined by an appearance probability (AP) matrix; and
- an element AP_{ij} of the AP matrix is determined by below Equations,

$$AP_{ij} = \frac{WF_{ij}}{\sum_{i,j} WF_{ij}}$$

$$WF_{ab} = \sum_{\substack{\text{For all protein pairs } p_i, q_j \\ \text{s.t. } \langle a, b \rangle \in dc\text{-pair}(p_i, q_j)}} \frac{1}{|dc(p_i)| \times |dc(q_j)|}$$

4. The prediction method as set forth in claim 1, wherein the step (c) includes the sub-steps of:

- (c-1) obtaining probability values and distributions of the probability values by

applying the determined probability equation to the interacting and non-interacting sets of protein pairs;

(c-2) obtaining a probability value by applying the determined probability equation to the given protein pair; and

5 (c-3) computing a probability for determining which set the given protein pair belongs to based on the distributions of the interacting and non-interacting sets of protein pairs.

5. A system for predicting the interaction between proteins comprising:

10 a probability data storing unit for obtaining appearance frequency information of a designated domain combination selected from each of interacting and non-interacting sets of protein pairs and storing the obtained appearance frequency information;

a probability equation determining unit for determining a probability equation applied to predict the interaction between two proteins, randomly selected, using the stored appearance frequency information of domain combination; and

15 a probability equation computing unit for computing an interaction probability between the two proteins using the determined probability equation.

6. The prediction system as set forth in claim 5,

20 wherein the probability equation computing unit obtains probability values and distributions of the probability values by applying the determined probability equation to the interacting and non-interacting sets of protein pairs, obtains a probability value by applying the determined probability equation to the given protein pair, and then computes a probability for determining which set the protein pair belongs to based on the distributions of the interacting and non-interacting sets of protein pairs.

25

7. A computer readable recording media for recording a program for achieving, by means of a computer, the prediction method as set forth in any one of claims 1 to 4.